An impossibility theorem for paired comparisons

László Csató*

Institute for Computer Science and Control, Hungarian Academy of Sciences (MTA SZTAKI)
Laboratory on Engineering and Management Intelligence, Research Group of Operations Research and Decision Systems

Corvinus University of Budapest (BCE)
Department of Operations Research and Actuarial Sciences

Budapest, Hungary

December 2, 2016

Abstract

In several decision-making problems, alternatives should be ranked on the basis of paired comparisons between them. An axiomatic approach for the universal ranking problem with arbitrary preference intensities, incomplete and multiple comparisons is presented. In particular, two basic properties – independence of irrelevant matches and self-consistency – are considered. It is revealed that there exists no ranking method satisfying both requirements at the same time. The impossibility result holds under various restrictions on the set of ranking problems, however, it does not emerge in the case of round-robin tournaments. An interesting and more general possibility result is obtained by weakening independence of irrelevant matches through the concept of macrovertex.

***JEL* classification number**: C44, D71

***AMS* classification number**: 15A06, 91B14

Keywords: Preference aggregation; paired comparison; tournament ranking; axiomatic approach; impossibility

1 Introduction

Paired-comparison based ranking emerges in many fields of science such as social choice theory (Chebotarev and Shamis, 1998), sports (Landau, 1895, 1914; Zermelo, 1929), or psychology (Thurstone, 1927). Here a general version of the problem, allowing for different preference intensities (including ties) as well as incomplete and multiple comparisons between two objects, is addressed.

The paper contributes to this field by the formulation of an impossibility theorem: it turns out that two axioms, independence of irrelevant matches – used, among others,

* e-mail: laszlo.csato@uni-corvinus.hu

in characterizations of Borda ranking by Rubinstein (1980) and Nitzan and Rubinstein (1981) and recently discussed by González-Díaz et al. (2014) – and self-consistency – a less known but intuitive property, introduced in Chebotarev and Shamis (1997) – cannot be satisfied at the same time. We also investigate domain restrictions and the weakening of the properties in order to get some positive results.

Our main theorem reinforces that while the row sum (sometimes called Borda or score) ranking has favourable properties in the case of round-robin tournaments, its application can be attacked when there are some complications like incomplete comparisons. A basket case is a Swiss-system tournament, where players with weaker opponents can score the same number of points more easily, so row sum seems to be a bad choice despite it is widely used (see Csató (2016) for details).

The current paper can be regarded as a supplement to the findings of previous axiomatic discussions in the field (Altman and Tennenholtz, 2008; Chebotarev and Shamis, 1998; González-Díaz et al., 2014) by highlighting some unknown connections among axioms. Furthermore, our impossibility result gives mathematical justification for a comment appearing in the axiomatic analysis of scoring procedures by González-Díaz et al. (2014): 'when players have different opponents (or face opponents with different intensities), IIM[1] is a property one would rather not have' (p. 165). The strength of this axiom is clearly showed by the main theorem.

The study is structured as follows. Section 2 presents the setting of the ranking problem, and the definitions of ranking methods examined. In Section 3, two axioms are evoked in order to present a significant impossibility result. Section 4 follows different ways to achieve possibility. Finally, some concluding remarks are given in Section 5.

2 Preliminaries

The following part of the paper discusses the representation of ranking problems and introduces some scoring procedures investigated later.

2.1 The ranking problem

Let $N = \{X_1, X_2, \dots, X_n\}$, $n \in \mathbb{N}$ be the *set of objects* and $T = [t_{ij}] \in \mathbb{R}^{n \times n}$ be a *tournament matrix* such that $t_{ij} + t_{ji} \in \mathbb{N}$. t_{ij} represents the aggregated score of object X_i against X_j, $t_{ij}/(t_{ij} + t_{ji})$ can be interpreted as the likelihood that object X_i is better than object X_j. $t_{ii} = 0$ is assumed for all $X_i \in N$. Possible derivations of the tournament matrix can be found in González-Díaz et al. (2014) and Csató (2015).

The pair (N, T) is called a *ranking problem*. The set of ranking problems with $|N| = n$ is denoted by $\mathcal{R}^n$.

A *scoring procedure* f is an $\mathcal{R}^n \to \mathbb{R}^n$ function giving a rating for each object. Any scoring method immediately determines a ranking (a transitive and complete weak order on the set $N \times N$) $\succeq$ by $f_i(N, T) \geq f_j(N, T)$ meaning that X_i is ranked weakly above X_j, denoted by $X_i \succeq X_j$. Every scoring method can be considered as a *ranking method*. This paper discusses only ranking methods derived from scoring procedures.

A ranking problem (N, T) is associated with the skew-symmetric *results matrix* $R = T - T^\top = [r_{ij}] \in \mathbb{R}^{n \times n}$ and the symmetric *matches matrix* $M = T + T^\top = [m_{ij}] \in \mathbb{N}^{n \times n}$

[1] Independence of irrelevant matches, discussed in Section 3.1.

such that m_{ij} is the number of the comparisons between X_i and X_j, whose outcome is given by r_{ij}. Matrices R and M also determine the tournament matrix by $T = (R+M)/2$.

Remark 2.1. Any ranking problem $(N, T) \in \mathcal{R}^n$ can be denoted analogously as (N, R, M) with the restriction $|r_{ij}| \leq m_{ij}$ for all $X_i, X_j \in N$, that is, the outcome of any paired comparison between two objects cannot 'exceed' their number of matches.

Despite the description through results and matches matrices is not parsimonious, usually the notation (N, R, M) will be used because it helps to define certain axioms.

The class of these universal ranking problems has some meaningful subsets.

Definition 2.1. *Special classes of ranking problems*: A ranking problem $(N, R, M) \in \mathcal{R}^n$ is called

- *balanced* if $\sum_{X_k \in N} m_{ik} = \sum_{X_k \in N} m_{jk}$ for all $X_i, X_j \in N$;
- *round-robin* if $m_{ij} = m_{k\ell}$ for all $X_i \neq X_j$ and $X_k \neq X_\ell$;
- *unweighted* if $m_{ij} \in \{0; 1\}$ for all $X_i, X_j \in N$;
- *extremal* if $|r_{ij}| \in \{0; m_{ij}\}$ for all $X_i, X_j \in N$.

The first three subsets pose restrictions on the matches matrix M. In a balanced ranking problem, all objects should have the same number of comparisons – a typical example is a Swiss-system tournaments (provided the number of participants is even). In a round-robin ranking problem, the number of comparisons between any pair of objects is the same – a typical example (of double round-robin) can be the qualification for soccer tournaments like UEFA European Championship. It does not allows for incomplete comparisons. Finally, in an unweighted ranking problem, multiple comparisons may not occur.

Extremal ranking problems mean conditions on the results matrix: the outcome of a comparison can only be a complete win ($r_{ij} = m_{ij}$), a draw ($r_{ij} = 0$), or a maximal loss ($r_{ij} = -m_{ij}$). In other words, preferences have no intensity, however, ties are allowed.

One can also consider any intersection of these special classes.

Notation 2.1. The set of balanced ranking problems is denoted by $\mathcal{R}_B$.
The set of round-robin ranking problems is denoted by $\mathcal{R}_R$.
The set of unweighted ranking problems is denoted by $\mathcal{R}_U$.
The set of extremal ranking problems is denoted by $\mathcal{R}_E$.

Remark 2.2. A round-robin ranking problem is balanced: $\mathcal{R}_R \subset \mathcal{R}_B$.

The *number of comparisons* of object X_i is $d_i = \sum_{X_j \in N} m_{ij}$ for all $X_i \in N$ and the *maximal number of comparisons* is $m = \max_{X_i, X_j \in N} m_{ij}$.

Remark 2.3. A ranking problem is balanced if and only if $d_i = d$ for all $X_i \in N$.
A ranking problem is round-robin if and only if $m_{ij} = m$ for all $X_i, X_j \in N$.
A ranking problem is unweighted if and only if $m = 1$.[2]

Matrix M can be represented by an undirected multigraph $G := (V, E)$, where the vertex set V corresponds to the object set N, and the number of edges between objects X_i and X_j is equal to m_{ij}. This graph G is said to be the *comparison multigraph* of the ranking problem (N, R, M), it is independent of the results matrix R. The *Laplacian matrix* $L = [\ell_{ij}] \in \mathbb{R}^{n \times n}$ of graph G is given by $\ell_{ij} = -m_{ij}$ for all $X_i \neq X_j$ and $\ell_{ii} = d_i$ for all $X_i \in N$.

[2] While $m_{ij} \in \{0; 1\}$ for all $X_i, X_j \in N$ allows for $m = 0$, it leads to a meaningless ranking problem without any comparison.

Remark 2.4. The degree of node X_i in the comparison multigraph is d_i.

A ranking problem $(N, R, M) \in \mathcal{R}^n$ is called *connected* or *unconnected* if its comparison multigraph is connected or unconnected, respectively.

2.2 Some ranking methods

In the following, some scoring procedures are presented, which associate an n-dimensional vector with any ranking problem $(N, R, M) \in \mathcal{R}^n$. They will be used only for ranking purposes, so they can be called ranking methods.

Let $\mathbf{e} \in \mathbb{R}^n$ denote the column vector with $e_i = 1$ for all $i = 1, 2, \dots, n$. Let $I \in \mathbb{R}^{n \times n}$ be the identity matrix.

The first scoring method does not take the comparison structure into account.

Definition 2.2. *Row sum*: $\mathbf{s}(N, R, M) = R\mathbf{e}$.

The following *parametric* procedure has been constructed axiomatically by Chebotarev (1989) and thoroughly analysed in Chebotarev (1994).

Definition 2.3. *Generalized row sum*: it is the unique solution $\mathbf{x}(\varepsilon)(N, R, M)$ of the system of linear equations $(I + \varepsilon L)\mathbf{x}(\varepsilon)(N, R, M) = (1 + \varepsilon mn)\mathbf{s}(N, R, M)$, where $\varepsilon > 0$ is a parameter.

Generalized row sum adjusts the row sum s_i by accounting for the performance of objects compared with X_i, and adds an infinite depth to this argument: the row sums of all objects available on a path from X_i appear in the calculation. ε indicates the importance attributed to this correction.

Remark 2.5. Generalized row sum results in row sum if $\varepsilon \to 0$: $\lim_{\varepsilon \to 0} \mathbf{x}(\varepsilon)(N, R, M) = \mathbf{s}(N, R, M)$.

The row sum and generalized row sum rankings are unique and easily computable from a system of linear equations for all ranking problems $(N, R, M) \in \mathcal{R}^n$.

The least squares method was suggested by Thurstone (1927) and Horst (1932).

Definition 2.4. *Least squares*: it is the solution $\mathbf{q}(N, R, M)$ of the system of linear equations $L\mathbf{q}(N, R, M) = \mathbf{s}(N, R, M)$ and $\mathbf{e}^\top \mathbf{q}(N, R, M) = 0$.

Remark 2.6. Generalized row sum ranking coincides with least squares ranking if $\varepsilon \to \infty$: $\lim_{\varepsilon \to \infty} \mathbf{x}(\varepsilon)(N, R, M) = mn\mathbf{q}(N, R, M)$.

The least squares ranking is unique if and only if the ranking problem $(N, R, M) \in \mathcal{R}^n$ is connected (Chebotarev and Shamis, 1999, p. 220). The ranking of unconnected objects makes no sense. Nonetheless, the least squares ranking can be made unique for all ranking problems $(N, R, M) \in \mathcal{R}^n$ if definition 2.4 is applied for all ranking subproblems with a connected comparison multigraph.

An extensive analysis and a graph interpretation of the least squares method as well as further references can be found in Csató (2015).

3 The impossibility result

In this section, a natural axiom of independence and a kind of monotonicity property is recalled. Our main result illustrates the impossibility of satisfying the two requirements simultaneously.

3.1 Independence of irrelevant matches

The following property appears as *independence* in Rubinstein (1980, Axiom III) and Nitzan and Rubinstein (1981, Axiom 5) in the case of round-robin ranking problems. The name independence of irrelevant matches has been used by González-Díaz et al. (2014) on the domain of all ranking problems. It deals with the effects of certain changes in the tournament matrix.

Axiom 3.1. *Independence of irrelevant matches* (IIM): Let $(N,T),(N,T') \in \mathcal{R}^n$ be two ranking problems and $X_i, X_j, X_k, X_\ell \in N$ be four different objects such that (N,T) and (N,T') are identical but $t'_{k\ell} \neq t_{k\ell}$. Scoring procedure $f : \mathcal{R}^n \to \mathbb{R}^n$ is called *independent of irrelevant matches* if $f_i(N,T) \geq f_j(N,T) \Rightarrow f_i(N,T') \geq f_j(N,T')$.

IIM means that 'remote' comparisons – not involving objects X_i and X_j – do not affect the relative ranking of X_i and X_j. Changing the matches matrix may lead to an unconnected ranking problem.

Remark 3.1. Property IIM has a meaning if $n \geq 4$.

Sequential application of independence of irrelevant matches can lead to any ranking problem $(N,\bar{T}) \in \mathcal{R}^n$, for which $\bar{t}_{gh} = t_{gh}$ if $\{X_g, X_h\} \cap \{X_i, X_j\} \neq \emptyset$, but all other paired comparisons are arbitrary.

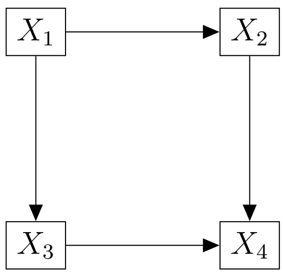
Lemma 3.1. *The row sum method is independent of irrelevant matches.*

Proof. It follows from Definition 2.2. □

3.2 Self-consistency

The next axiom, introduced by (Chebotarev and Shamis, 1997), may require some explanation. It is motivated with an example using the language of preference aggregation.

Figure 1: Ranking problem of Example 3.1



Example 3.1. Consider the ranking problem $(N,R,M) \in \mathcal{R}^4_B \cap \mathcal{R}^4_U \cap \mathcal{R}^4_E$ with results and matches matrices

$$R = \begin{bmatrix} 0 & 1 & 1 & 0 \\ -1 & 0 & 0 & 1 \\ -1 & 0 & 0 & 1 \\ 0 & -1 & -1 & 0 \end{bmatrix} \text{ and } M = \begin{bmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{bmatrix}.$$

It is shown in Figure 1: a directed edge from node X_i to X_j indicates a complete win of X_i over X_j (and a complete loss of X_j against X_i). This correspondence will be used in further examples, too.

The situation in Example 3.1 can be interpreted as follows. A voter prefers alternative X_1 to X_2 and X_3, but says nothing about X_4. Another voter prefers X_2 to X_3 and X_4, but has no opinion on X_1.

Despite it is difficult to make a good decision on the basis of such incomplete preferences, sometimes it cannot be avoided. It leads to the question, which principles should be followed by the aggregation procedure. It seems reasonable that X_i should be judged better than X_j if one of the following holds:

✠1 X_i achieves better results against the same opponents;

✠2 X_i achieves better results against opponents with the same strength;

✠3 X_i achieves the same results against stronger opponents;

✠4 X_i achieves better results against stronger opponents.

Furthermore, X_i should have the same rank as X_j if one of the following holds:

✠5 X_i achieves the same results against the same opponents;

✠6 X_i achieves the same results against opponents with the same strength.

In order to apply these principles, one should measure the strength of opponents. It is provided by the scoring method itself, hence the name of this axiom is *self-consistency*. Consequently, condition ✠1 is a special case of condition ✠2 (the same opponents have naturally the same strength) as well as condition ✠5 is implied by condition ✠6.

What does self-consistency mean in Example 3.1? First, $X_2 \sim X_3$ since condition ✠5. Second, $X_1 \succ X_4$ should hold since condition ✠1 as $r_{12} > r_{42}$ and $r_{13} > r_{43}$. The requirements above can also be applied for objects which have different opponents. Assume that $X_1 \preceq X_2$. Then condition ✠4 results in $X_1 \succ X_2$ because of $X_2 \succeq X_1$, $r_{12} > r_{21}$ and $X_3 \sim X_2 \succeq X_1 \succ X_4$, $r_{13} = r_{24}$. It is a contradiction, therefore $X_1 \succ (X_2 \sim X_3)$. Similarly, assume that $X_2 \preceq X_4$. Then condition ✠4 results in $X_2 \succ X_4$ because of $X_1 \succ X_3$ (derived above), $r_{21} = r_{43}$ and $X_4 \succeq X_2 \sim X_3$, $r_{24} > r_{43}$. It is a contradiction, therefore $(X_2 \sim X_3) \succ X_4$. To summarize, only the ranking $X_1 \succ (X_2 \sim X_3) \succ X_4$ is allowed by self-consistency.

The above requirement can be formalized in the following way.

Definition 3.1. *Opponent set*: Let $(N, R, M) \in \mathcal{R}_U^n$ be an unweighted ranking problem. The *opponent set* of object X_i is $O_i = \{X_j : m_{ij} = 1\}$

Objects of the opponent set O_i are called *opponents* of X_i.

Remark 3.2. $|O_i| = |O_j|$ for all $X_i, X_j \in N$ if and only if the ranking problem is balanced.

Notation 3.1. Let $(N, R, M) \in \mathcal{R}_U^n$ be an unweighted ranking problem, $X_i, X_j \in N$ be two different objects and $g : O_i \leftrightarrow O_j$ be a one-to-one correspondence between the opponents of X_i and X_j. Then $\mathfrak{g}$ is given by $X_{\mathfrak{g}(k)} = g(X_k)$.

In order to make judgements like an object has stronger opponents, at least a partial order on the opponent sets should be introduced.

Definition 3.2. *Partial order of opponent sets*: Let $(N, R, M) \in \mathcal{R}^n$ be a ranking problem and $f : \mathcal{R}^n \to \mathbb{R}^n$ be a scoring procedure. Opponents of X_i are at least as strong as opponents of X_j, denoted by $O_i \succeq O_j$, if there exists a one-to-one correspondence $g : O_i \leftrightarrow O_j$ such that $f_k(N, R, M) \geq f_{\mathfrak{g}(k)}(N, R, M)$ for all $X_k \in O_i$.

For instance, $O_1 \sim O_4$ and $O_2 \sim O_3$ in Example 3.1, while O_1 and O_2 are not comparable.

Not that conditions ✠1-✠6 could not imply $X_i \succeq X_j$ if $O_i \succ O_j$: an object with a weaker opponent set cannot be judged better.

Opponent sets has been defined onyl for unweighted ranking problems, but self-consistency can be applies for objects which have the same number of comparisons. The extension is achieved by decomposition of ranking problems.

Definition 3.3. *Sum of ranking problems*: Let $(N, R, M), (N, R', M') \in \mathcal{R}^n$ be two ranking problems with the same object set N. The *sum* of these ranking problems is the ranking problem $(N, R + R', M + M') \in \mathcal{R}^n$.

Summing of ranking problems may have a natural interpretation. For example, they can contain the preferences of voters in two cities of the same country, or the paired comparisons of players in the first and second half of the year.

Definition 3.3 implies that any ranking problem can be decomposed into unweighted ranking problems, that is, it can be obtained as a sum of unweighted ranking problems. However, while the sum of ranking problems is unique, a ranking problem may have a number of possible decompositions.

Axiom 3.2. *Self-consistency* (*SC*) (Chebotarev and Shamis, 1997): Let $(N, R, M) \in \mathcal{R}^n$ be a ranking problem such that $R = \sum_{p=1}^{m} R^{(p)}$, $M = \sum_{p=1}^{m} M^{(p)}$ and $(N, R^{(p)}, M^{(p)}) \in \mathcal{R}_U^n$ is an unweighted ranking problem for all $p = 1, 2, \dots, m$. Let $X_i, X_j \in N$ be two objects and $f : \mathcal{R}^n \to \mathbb{R}^n$ be a scoring procedure such that for all $p = 1, 2, \dots, m$ there exists a one-to-one mapping g from $O_i^{(p)}$ onto $O_j^{(p)}$, where $r_{ik}^{(p)} \geq r_{j\mathfrak{g}(k)}^{(p)}$ and $f_k(N, R, M) \geq f_{\mathfrak{g}(k)}(N, R, M)$.
f is called *self-consistent* if $f_i(N, R, M) \geq f_j(N, R, M)$, furthermore, $f_i(N, R, M) > f_j(N, R, M)$ if at least one of the above inequalities is strict.

Self-consistency provides that if object X_i is obviously not worse than object X_j, then it is not ranked lower, furthermore, if it is better, then it is ranked higher. In other words, it formalizes conditions ✠1-✠6. Note that self-consistency can also be interpreted as a property of a ranking.

The application of self-consistency is not trivial because of the various opportunities for decomposition into unweighted ranking problems.

Remark 3.3. Self-consistency restricts the relative ranking of objects X_i and X_j only if $d_i = d_j$ since there should exist a one-to-one mapping between $O_i^{(p)}$ and $O_j^{(p)}$ for all $p = 1, 2, \dots, m$.

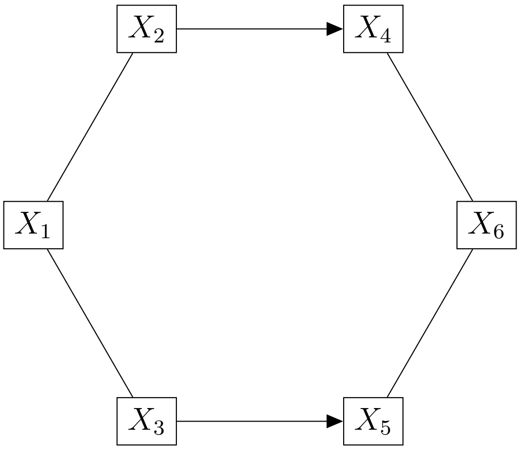
Remark 3.3 immediately means that *SC* does not fully characterize a ranking. It is not true even on the set of balanced ranking problems as the following example shows.

Lemma 3.2. *Self-consistency does not fully characterize a ranking method on the set of balanced, unweighted and extremal ranking problems $\mathcal{R}_B \cap \mathcal{R}_U \cap \mathcal{R}_E$.*

Proof. The statement can be verified by an example where at least two rankings are allowed by *SC*.

Example 3.2. Consider the ranking problem $(N, R, M) \in \mathcal{R}_B^6 \cap \mathcal{R}_U^6 \cap \mathcal{R}_E^6$. It is shown in Figure 2: a directed edge from node X_i to X_j indicates a complete win of X_i over X_j in one comparison (as in Example 3.1) and an undirected edge from node X_i to X_j represents a draw in one comparison between the two objects.

Figure 2: Ranking problem of Example 3.2



Consider the ranking $\succeq^1$ such that $(X_1 \sim^1 X_2 \sim^1 X_3) \succ^1 (X_4 \sim^1 X_5 \sim^1 X_6)$. The opponent sets are $O_1 = \{X_2, X_3\}$, $O_2 = \{X_1, X_4\}$, $O_3 = \{X_1, X_5\}$, $O_4 = \{X_2, X_6\}$, $O_5 = \{X_3, X_6\}$ and $O_6 = \{X_4, X_5\}$, so $O_1 \succ (O_2 \sim O_3 \sim O_4 \sim O_5) \succ O_6$. The results of X_2 and X_3 are $(0; 1)$, the results of X_1 and X_6 are $(0; 0)$ and the results of X_4 and X_5 are $(-1; 0)$. For objects with the same results, SC implies $X_2 \sim X_3$, $X_4 \sim X_5$ and $X_1 \succ X_6$ (conditions ✠3 and ✠6), which hold in $\succeq^1$. For objects with the different results, SC leads to $X_1 \succ X_4$, $X_2 \succ X_4$ and $X_2 \succ X_6$ taking the strength of opponents into account (condition ✠2). These requirements are also met by the ranking $\succeq^1$. Self-consistency imposes no other restrictions, therefore the ranking $\succeq^1$ satisfies this axiom.

Now consider the ranking $\succeq^2$ such that $X_1 \prec^2 (X_2 \sim^2 X_3) \prec^2 (X_4 \sim^2 X_5) \prec^2 X_6$. The opponent sets are unchanged, but their partial order is different, given by $O_1 \prec (O_4 \sim O_5)$, $O_1 \prec O_6$, $(O_2 \sim O_3) \prec (O_4 \sim O_5)$ and $(O_2 \sim O_3) \prec O_6$ (the opponents of X_1 and X_2 as well as X_4 and X_6 cannot be compared). For objects with the same results, SC implies $X_2 \sim X_3$, $X_4 \sim X_5$ and $X_1 \prec X_6$ (conditions ✠3 and ✠6), which hold in $\succeq^2$. For objects with different results, SC leads to $X_1 \prec X_2$ taking the strength of opponents into account (condition ✠2). This requirement is also met by the ranking $\succeq^2$. Self-consistency imposes no other restrictions, therefore the ranking $\succeq^2$ satisfies this axiom, too.

To conclude, rankings $\succeq^1$ and $\succeq^2$ are self-consistent.[3] □

Lemma 3.3. *The generalized row sum and least squares methods are self-consistent.*

Proof. See Chebotarev and Shamis (1998, Theorem 5). □

3.3 The connection of independence of irrelevant matches and self-consistency

So far we have presented two axioms, IIM and SC. It turns out that they cannot be satisfied at the same time.

Theorem 3.1. *There exists no scoring procedure that is independent of irrelevant matches and self-consistent.*

Proof. The contradiction of the two properties is proved by the following example.

[3] The ranking opposite to $\succeq^2$ also satisfies the axiom SC.

Figure 3: Ranking problems of Example 3.3

(a) Ranking problem (N, R, M) (b) Ranking problem (N, R', M)

Example 3.3. Let $(N, R, M), (N, R', M) \in \mathcal{R}_B^4 \cap \mathcal{R}_U^4 \cap \mathcal{R}_E^4$ be the ranking problems in Figure 3 with the results and matches matrices

$$R = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \end{bmatrix}, \; R' = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{bmatrix}, \text{ and } M = \begin{bmatrix} 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \end{bmatrix}.$$

The opponent sets are $O_1 = O_4 = \{X_2, X_3\}$ and $O_2 = O_3 = \{X_1, X_4\}$ in both ranking problems. Assume to the contrary that there exists a scoring procedure f, which is independent of irrelevant matches and self-consistent. IIM means that $f_1(N, R, M) \geq f_2(N, R, M) \iff f_1(N, R', M) \geq f_2(N, R', M)$.

a) Consider the (identity) one-to-one mapping $g_{14} : O_1 \leftrightarrow O_4$, where $g_{14}(X_2) = X_2$ and $g_{14}(X_3) = X_3$. Since $r_{12} = r_{42} = 0$ and $0 = r_{13} > r_{43} = -1$, g_{14} satisfies condition ✠1 of SC, hence $f_1(N, R, M) > f_4(N, R, M)$.

b) Consider the identity one-to-one mapping $g_{32} : O_3 \leftrightarrow O_2$, where $g_{32}(X_1) = X_1$ and $g_{32}(X_4) = X_4$. Since $r_{31} = r_{21} = 0$ and $1 = r_{34} > r_{24} = 0$, g_{32} satisfies condition ✠1 of SC, hence $f_3(N, R, M) > f_2(N, R, M)$.

c) Suppose that $f_2(N, R, M) \geq f_1(N, R, M)$, implying $f_3(N, R, M) > f_4(N, R, M)$. Consider the one-to-one-correspondence $g_{12} : O_1 \leftrightarrow O_2$, where $g_{12}(X_2) = X_1$ and $g_{12}(X_3) = X_4$. Since $r_{12} = r_{21} = 0$ and $r_{13} = r_{24} = 0$, g_{12} satisfies condition ✠3 of SC, hence $f_1(N, R, M) > f_2(N, R, M)$. It is a contradiction.

Thus only $f_1(N, R, M) > f_2(N, R, M)$ is allowed.

Note that ranking problem (N, R', M) can be obtained from (N, R, M) by the permutation $\sigma : N \to N$ such that $\sigma(X_1) = X_2$, $\sigma(X_2) = X_1$, $\sigma(X_3) = X_4$ and $\sigma(X_4) = X_3$. The above argument results in $f_2(N, R', M) > f_1(N, R', M)$, contrary to independence of irrelevant matches.

To conclude, no scoring procedure can meet IIM and SC simultaneously. □

Corollary 3.1. *The score method may violate self-consistency.*

Proof. It is an immediate consequence of Lemma 3.1 and Theorem 3.1.[4] □

[4] Chebotarev and Shamis (1998, Theorem 5) erroneously states that the score method trivially satisfies self-consistency.

Corollary 3.2. *The generalized row sum and least squares methods may violate independence of irrelevant matches.*

Proof. It follows from Lemma 3.3 and Theorem 3.1. □

Corollary 3.3. *IIM and SC are logically independent axioms.*

Proof. It is the consequence of Corollaries 3.1 and 3.2. □

4 How to achieve possibility?

Impossibility results, like the one in Theorem 3.1, can be avoided in at least two ways: by restrictions on the class of ranking problems considered, or by the weakening of one or more axioms.

4.1 Domain restrictions

Besides the natural subclasses of ranking problems introduced in Definition 2.1, the number of objects can be limited, too.

Lemma 4.1. *The generalized row sum and least squares methods are independent of irrelevant matches and self-consistent on the set of ranking problems with at most three objects $\mathcal{R}^n | n \leq 3$.*

Proof. According to Remark 3.1, IIM means nothing on the set $\mathcal{R}^n | n \leq 3$, so any self-consistent scoring procedure is appropriate, thus Lemma 3.3 provides the result. □

However, if four objects are allowed, the situation is much severe.

Lemma 4.2. *There exists no scoring procedure $f : \mathcal{R}^n \to \mathbb{R}^n$ that is independent of irrelevant matches and self-consistent on the set of balanced, unweighted and extremal ranking problems with four objects $\mathcal{R}^4_B \cap \mathcal{R}^4_U \cap \mathcal{R}^4_E$.*

Proof. The ranking problems of Example 3.3, used for verifying the impossibility in Theorem 3.1, are from the set $\mathcal{R}^4_B \cap \mathcal{R}^4_U \cap \mathcal{R}^4_E$. □

Lemma 4.2 does not deal with the fourth subset in Definition 2.1, the class of round-robin ranking problems. Then another possibility result emerges.

Proposition 4.1. *The row sum, generalized row sum and least squares methods are independent of irrelevant matches and self-consistent on the set of round-robin ranking problems $\mathcal{R}_R$.*

Proof. Due to axioms *agreement* (Chebotarev, 1994, Property 3) and *score consistency* (González-Díaz et al., 2014), the generalized row sum and least squares ranking methods coincide with the row sum on the set of $\mathcal{R}_R$, so Lemmata 3.1 and 3.2 provide IIM and SC, respectively. □

Perhaps it is not by chance that characterizations of the row sum method were suggested on this – or even more restricted – domain (Young, 1974; Hansson and Sahlquist, 1976; Rubinstein, 1980; Nitzan and Rubinstein, 1981; Henriet, 1985; Bouyssou, 1992).

4.2 Weakening of independence of irrelevant matches

For the relaxation of IIM, a property discussed by Chebotarev (1994) is presented after a short introduction.

Definition 4.1. *Macrovertex* (Chebotarev, 1994, Definition 3.1): Let $(N, R, M) \in \mathcal{R}^n$ be a ranking problem. Object set $V \subseteq N$ is called *macrovertex* if $m_{ik} = m_{jk}$ for all $X_i, X_j \in V$ and $X_k \in N \setminus V$.

Objects in a macrovertex have the same number of comparisons against any object outside the macrovertex. The comparison structure in V and $N \setminus V$ can be arbitrary. Existence of a macrovertex depends only on the matches matrix M, or, in other words, on the comparison multigraph of the ranking problem.

Axiom 4.1. *Macrovertex independence (MVI)* (Chebotarev, 1994, Property 8): Let $V \subseteq N$ be a macrovertex in ranking problems $(N, T), (N, T') \in \mathcal{R}^n$ and $X_i, X_j \in V$ be two different objects such that (N, T) and (N, T') are identical but $t'_{ij} \neq t_{ij}$. Scoring procedure $f : \mathcal{R}^n \to \mathbb{R}^n$ is called *macrovertex independent* if $f_k(N, T) \geq f_\ell(N, T) \Rightarrow f_k(N, T') \geq f_\ell(N, T')$ for all $X_k, X_\ell \in N \setminus V$.

Macrovertex independence says that the relative ranking of objects outside a macrovertex is independent of the number and result of comparisons between the objects inside the macrovertex.

Corollary 4.1. *IIM implies MVI.*

Note that if V is a macrovertex, then $N \setminus V$ is not necessarily another macrovertex. Hence a pair of property MVI can be discussed.

Axiom 4.2. *Macrovertex autonomy (MVA)*: Let $V \subseteq N$ be a macrovertex in ranking problems $(N, T), (N, T') \in \mathcal{R}^n$ and $X_k, X_\ell \in N \setminus V$ be two different objects such that (N, T) and (N, T') are identical but $t'_{k\ell} \neq t_{k\ell}$. Scoring procedure $f : \mathcal{R}^n \to \mathbb{R}^n$ is called *macrovertex autonom* if $f_i(N, T) \geq f_j(N, T) \Rightarrow f_i(N, T') \geq f_j(N, T')$ for all $X_i, X_j \in V$.

Macrovertex autonomy says that the relative ranking of objects inside a macrovertex is independent of the number and result of comparisons between the objects outside the macrovertex.

Corollary 4.2. *IIM implies MVA.*

Similarly to IIM, changing the matches matrix – as allowed by properties MVI and MVA – may lead to an unconnected ranking problem.

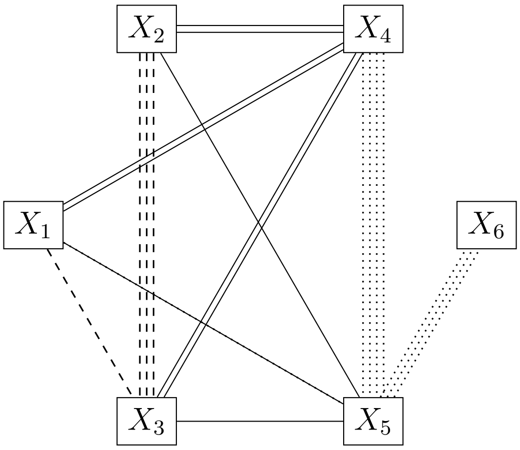
Example 4.1. Consider a ranking problem with the comparison multigraph in Figure 4. Object set $V = \{X_1, X_2, X_3\}$ is a macrovertex, the number of edges from any node inside V to any node outside V is the same (two to X_4, one to X_5, and zero to X_6). V remains a macrovertex if comparisons inside V (represented by dashed edges) or comparisons outside V (dotted edges) are changed.

Macrovertex independence requires that the relative ranking of X_4, X_5 and X_6 does not depend on the number and result of comparisons between X_1, X_2 and X_3.

Macrovertex autonomy requires that the relative ranking of X_1, X_2 and X_3 does not depend on the number and result of comparisons between X_4, X_5 and X_6

The implications of MVI and MVA are clearly different since object set $N \setminus V = \{X_4, X_5, X_6\}$ is not a macrovertex because $m_{14} = 2 \neq 1 = m_{15}$.

Figure 4: Comparison multigraph of Example 4.1



Corollary 4.3. *The row sum method is macrovertex independent and macrovertex autonom.*

Proof. It is an immediate consequence of Lemma 3.1 and Corollaries 4.1 and 4.2. □

Lemma 4.3. *The generalized row sum and least squares methods are macrovertex independent and macrovertex autonom.*

Proof. Chebotarev (1994, Property 8) has shown that generalized row sum satisfies MVI. The proof remains valid in the limit $\varepsilon \to \infty$ if the least squares ranking is defined to be unique, that is, the sum of ratings of objects in all components of the comparison multigraph is zero.

Consider MVA. Let $\mathbf{s} = \mathbf{s}(N,T)$, $\mathbf{s}' = \mathbf{s}(N,T')$, $\mathbf{x} = \mathbf{x}(\varepsilon)(N,T)$, $\mathbf{x}' = \mathbf{x}(\varepsilon)(N,T')$ and $\mathbf{q} = \mathbf{q}(N,T)$, $\mathbf{q}' = \mathbf{q}(N,T')$. Let V be a macrovertex and $X_i, X_j \in V$ be two arbitrary objects. Suppose to the contrary that $x_i \geq x_j$, but $x'_i < x'_j$, hence $x'_i - x_i < x'_j - x_j$. Let $x'_k - x_k = \max_{X_g \in V}(x'_g - x_g)$ and $x'_\ell - x_\ell = \min_{X_g \in V}(x'_g - x_g)$, therefore $x'_k - x_k > x'_\ell - x_\ell$ and $x'_k - x_k \geq x'_g - x_g \geq x'_\ell - x_\ell$ for any object $X_g \in V$.

For object X_k, definition 2.3 results in

$$x_k = (1 + \varepsilon mn)s_k + \varepsilon \sum_{X_g \in V} m_{kg}(x_g - x_k) + \varepsilon \sum_{X_h \in N \setminus V} m_{kh}(x_h - x_k). \tag{1}$$

Apply (1) for object X_ℓ. The difference of these two equations is

$$\begin{aligned} x_k - x_\ell &= (1 + \varepsilon mn)(s_k - s_\ell) + \varepsilon \sum_{X_g \in V} \left[m_{kg}(x_g - x_k) - m_{\ell g}(x_g - x_\ell) \right] + \\ &\quad + \varepsilon \sum_{X_h \in N \setminus V} \left[m_{kh}(x_h - x_k) - m_{\ell h}(x_h - x_\ell) \right]. \end{aligned} \tag{2}$$

Note that $m_{kh} = m_{\ell h}$ for all $X_h \in N \setminus V$ since V is a macrovertex, therefore (2) is equivalent to

$$\begin{aligned} \left(1 + \varepsilon \sum_{X_h \in N \setminus V} m_{kh} \right)(x_k - x_\ell) &= (1 + \varepsilon mn)(s_k - s_\ell) + \\ &\quad + \varepsilon \sum_{X_g \in V} \left[m_{kg}(x_g - x_k) - m_{\ell g}(x_g - x_\ell) \right]. \end{aligned} \tag{3}$$

Apply (3) for the ranking problem (N, T'):

$$\left(1+\varepsilon \sum_{X_h \in N \setminus V} m'_{kh}\right)(x'_k - x'_\ell) = (1+\varepsilon mn)(s'_k - s'_\ell) + \\ +\varepsilon \sum_{X_g \in V} \left[m'_{kg}(x'_g - x'_k) - m'_{\ell g}(x'_g - x'_\ell)\right]. \quad (4)$$

Let $\Delta_{ij} = (x'_i - x'_j) - (x_i - x_j)$ for all $X_i, X_j \in V$. Note that $m'_{kh} = m_{kh}$ for all $X_h \in N \setminus V$, $m'_{kg} = m_{kg}$ and $m'_{\ell g} = m_{\ell g}$ for all $X_g \in V$ as well as $s'_k = s_k$ and $s'_\ell = s_\ell$ since only comparisons outside V may change. Take the difference of (4) and (3)

$$\left(1+\varepsilon \sum_{X_h \in N \setminus V} m_{kh}\right)\Delta_{k\ell} = \varepsilon \sum_{X_g \in V} \left(m_{kg}\Delta_{gk} - m_{\ell g}\Delta_{g\ell}\right). \quad (5)$$

Due to the choice of indices k and ℓ, $\Delta_{k\ell} > 0$ and $\Delta_{gk} \leq 0$, $\Delta_{g\ell} \geq 0$. It means that the left-hand side of (5) is positive, while its right-hand side is nonpositive, leading to a contradiction. Therefore only $x'_i - x_i = x'_j - x_j$, the condition required by MVA, can hold.

The same derivation can be implemented for the least squares method. With the notation $\Delta_{ij} = (q'_i - q'_j) - (q_i - q_j)$ for all $X_i, X_j \in V$, we get – analogously to (5) as $\varepsilon \to \infty$ –

$$\sum_{X_h \in N \setminus V} m_{kh}\Delta_{k\ell} = \sum_{X_g \in V} \left(m_{kg}\Delta_{gk} - m_{\ell g}\Delta_{g\ell}\right). \quad (6)$$

But $\Delta_{k\ell} > 0$, $\Delta_{gk} \leq 0$, and $\Delta_{g\ell} \geq 0$ is not enough for a contradiction now: (6) may hold if $\sum_{X_h \in N \setminus V} m_{kh} = 0$, namely, X_k is not connected to any object outside the macrovertex V as well as $\Delta_{gk} = 0$ and $\Delta_{g\ell} = 0$ when $m_{kg} = m_{\ell g} > 0$. However, if there exists no object $X_g \in N \setminus V$ such that $m_{kg} = m_{\ell g} > 0$, then there is no connection between object sets V and $N \setminus V$ since V is a macrovertex, and we have two independent ranking subproblems, where the least squares ranking is unique according to the extension of definition 2.4, so MVA holds. On the other hand, if there exists an object $X_g \in N \setminus V$ such that $m_{kg} = m_{\ell g} > 0$, then $\Delta_{gk} = 0$ and $\Delta_{g\ell} = 0$, but $\Delta_{k\ell} = \Delta_{g\ell} - \Delta_{gk} > 0$, which is a contradiction. Therefore $q'_i - q_i = q'_j - q_j$, the condition required by MVA, holds. □

Lemma 4.3 results in another possibility result.

Proposition 4.2. *The generalized row sum and least squares methods are macrovertex autonom, macrovertex independent and self-consistent.*

It turns out that this result is more general than the one obtained by restricting the domain to round-robin raking problems.

Remark 4.1. Let $(N, R, M) \in \mathcal{R}^n_R$ be a round-robin ranking problem. Object set $V \subseteq N$ is a macrovertex.

Corollary 4.4. *A macrovertex autonom or a macrovertex independent scoring procedure satisfies independence of irrelevant matches on the set of round-robin raking problems* $\mathcal{R}_R$.

Proof. Let $(N, T), (N, T') \in \mathcal{R}^n$ be two ranking problems and $X_i, X_j, X_k, X_\ell \in N$ be four different objects such that (N, T) and (N, T') are identical but $t'_{k\ell} \neq t_{k\ell}$.

Consider the macrovertex $V = \{X_i, X_j\}$. Macrovertex autonomy means $f_i(N, T) \geq f_j(N, T) \Rightarrow f_i(N, T') \geq f_j(N, T')$, the condition required by IIM.

Consider the macrovertex $V' = \{X_k, X_\ell\}$. Macrovertex independence means $f_i(N, T) \geq f_j(N, T) \Rightarrow f_i(N, T') \geq f_j(N, T')$, the condition required by IIM. □

4.3 Weakening of self-consistency

Despite its complicated form, self-consistency seems to be more difficult to debate than independence of irrelevant matches. There is an obvious way to soften this axiom by being more tolerant in the case of opponents: X_i should be judged at least as good as X_j if it achieves the same result against stronger opponents (instead of condition ✠3, which demands that X_i is strictly better than X_j).

Axiom 4.3. *Weak self-consistency* (WSC): Let $(N, R, M) \in \mathcal{R}^n$ be a ranking problem such that $R = \sum_{p=1}^{m} R^{(p)}$, $M = \sum_{p=1}^{m} M^{(p)}$ and $(N, R^{(p)}, M^{(p)}) \in \mathcal{R}^n_U$ is an unweighted ranking problem for all $p = 1, 2, \dots, m$. Let $X_i, X_j \in N$ be two objects and $f : \mathcal{R}^n \to \mathbb{R}^n$ be a scoring procedure such that for all $p = 1, 2, \dots, m$ there exists a one-to-one mapping g from $O_i^{(p)}$ onto $O_j^{(p)}$, where $r_{ik}^{(p)} \geq r_{j\mathfrak{g}(k)}^{(p)}$ and $f_k(N, R, M) \geq f_{\mathfrak{g}(k)}(N, R, M)$.
f is called *self-consistent* if $f_i(N, R, M) \geq f_j(N, R, M)$, furthermore, $f_i(N, R, M) > f_j(N, R, M)$ if $r_{ik}^{(p)} > r_{j\mathfrak{g}(k)}^{(p)}$ for at least one $p = 1, 2, \dots, m$.

Corollary 4.5. *SC implies WSC.*

Lemma 4.4. *The row sum method is weakly self-consistent.*

Proof. Let $(N, R, M) \in \mathcal{R}^n$ be a ranking problem such that $R = \sum_{p=1}^{m} R^{(p)}$, $M = \sum_{p=1}^{m} M^{(p)}$ and $(N, R^{(p)}, M^{(p)}) \in \mathcal{R}^n_U$ is an unweighted ranking problem for all $p = 1, 2, \dots, m$. Let $X_i, X_j \in N$ be two objects and $f : \mathcal{R}^n \to \mathbb{R}^n$ be a scoring procedure such that for all $p = 1, 2, \dots, m$ there exists a one-to-one mapping g from $O_i^{(p)}$ onto $O_j^{(p)}$, where $r_{ik}^{(p)} \geq r_{j\mathfrak{g}(k)}^{(p)}$ and $f_k(N, R, M) \geq f_{\mathfrak{g}(k)}(N, R, M)$.

It is clear that $s_i(N, R, M) = \sum_{p=1}^{m} \sum_{X_k \in O_i^{(p)}} r_{ik} \geq \sum_{p=1}^{m} \sum_{X_k \in O_j^{(p)}} r_{j\mathfrak{g}(k)} = s_j(N, R, M)$. Furthermore, $s_i(N, R, M) > s_j(N, R, M)$ if $r_{ik}^{(p)} > r_{j\mathfrak{g}(k)}^{(p)}$ for at least one $p = 1, 2, \dots, m$. □

According to Lemma 4.4, the violation of self-consistency by row sum (see Corollary 3.1) is a consequence of condition ✠3: the row sums of X_i and X_j are the same even if X_i achieves the same result as X_j against stronger opponents. It is an important theoretical argument against the use of row sum for ranking in tournaments organized in a not round-robin format, which supports the empirical findings of Csató (2016) in the case of Swiss-system chess team tournaments.

Proposition 4.3. *The row sum method is independent of irrelevant matches and weakly self-consistent.*

Proof. It follows from Lemmata 3.1 and 4.4. □

5 Conclusions

The current study has presented a basic impossibility result (Theorem 3.1) for ranking in a paired comparison-based setting, which allows for different preference intensities as well as incomplete and multiple comparisons. The theorem involves two independent axioms, one – called independence of irrelevant matches – posing a kind of independence concerning the relative ranking of two objects, and the other – self-consistency – requiring to rank objects with an obviously better performance higher.

We have also aspired to get some positive results. Domain restriction is fruitful in the case of round-robin tournaments (Proposition 4.1), whereas limiting the intensity and the number of preferences does not eliminate impossibility if the number of objects is meaningful (Lemma 4.2, but Lemma 4.1). Self-consistency has a natural weakening, satisfied by the trivial method of row sum besides IIM (Proposition 4.3), but it seems to be the more plausible property. Independence of irrelevant matches can be refined through the concept of macrovertex: the relative ranking of two objects should not depend on an outside comparison if the comparisons have a special structure. The implied possibility theorem (Proposition 4.2) is more general than the positive result in the case of of round-robin ranking problems (Proposition 4.1).

There remains an unexplored gap between our impossibility and possibility theorems since the latter allow for more than one scoring procedures. Actually, generalized row sum and least squares methods cannot be distinguished with respect to the properties examined.[5] The loss of independence of irrelevant matches makes characterizations in this general domain complicated since self-consistency is not an axiom easy to seize. Despite the challenges, axiomatic construction of scoring procedures means a natural continuation of the current research.

Acknowledgements

The research was supported by OTKA grant K 111797 and by the MTA Premium Post Doctorate Research Program.

References

Altman, A. and Tennenholtz, M. (2008). Axiomatic foundations for ranking systems. *Journal of Artificial Intelligence Research*, 31(1):473–495.

Bouyssou, D. (1992). Ranking methods based on valued preference relations: a characterization of the net flow method. *European Journal of Operational Research*, 60(1):61–67.

Chebotarev, P. (1989). Generalization of the row sum method for incomplete paired comparisons. *Automation and Remote Control*, 50(8):1103–1113.

Chebotarev, P. (1994). Aggregation of preferences by the generalized row sum method. *Mathematical Social Sciences*, 27(3):293–320.

Chebotarev, P. and Shamis, E. (1997). Constructing an objective function for aggregating incomplete preferences. In Tangian, A. and Gruber, J., editors, *Constructing Scalar-Valued Objective Functions*, volume 453 of *Lecture Notes in Economics and Mathematical Systems*, pages 100–124. Springer, Berlin-Heidelberg.

Chebotarev, P. and Shamis, E. (1998). Characterizations of scoring methods for preference aggregation. *Annals of Operations Research*, 80:299–332.

Chebotarev, P. and Shamis, E. (1999). Preference fusion when the number of alternatives exceeds two: indirect scoring procedures. *Journal of the Franklin Institute*, 336(2):205–226.

[5] Some of their differences are highlighted by González-Díaz et al. (2014).

Csató, L. (2015). A graph interpretation of the least squares ranking method. *Social Choice and Welfare*, 44(1):51–69.

Csató, L. (2016). On the ranking of a Swiss system chess team tournament. http://arxiv.org/abs/1507.05045.

González-Díaz, J., Hendrickx, R., and Lohmann, E. (2014). Paired comparisons analysis: an axiomatic approach to ranking methods. *Social Choice and Welfare*, 42(1):139–169.

Hansson, B. and Sahlquist, H. (1976). A proof technique for social choice with variable electorate. *Journal of Economic Theory*, 13(2):193–200.

Henriet, D. (1985). The Copeland choice function: an axiomatic characterization. *Social Choice and Welfare*, 2(1):49–63.

Horst, P. (1932). A method for determining the absolute affective value of a series of stimulus situations. *Journal of Educational Psychology*, 23(6):418–440.

Landau, E. (1895). Zur relativen Wertbemessung der Turnierresultate. *Deutsches Wochenschach*, 11:366–369.

Landau, E. (1914). Über Preisverteilung bei Spielturnieren. *Zeitschrift für Mathematik und Physik*, 63:192–202.

Nitzan, S. and Rubinstein, A. (1981). A further characterization of Borda ranking method. *Public Choice*, 36(1):153–158.

Rubinstein, A. (1980). Ranking the participants in a tournament. *SIAM Journal on Applied Mathematics*, 38(1):108–111.

Thurstone, L. L. (1927). A law of comparative judgment. *Psychological Review*, 34(4):273–286.

Young, H. P. (1974). An axiomatization of Borda's rule. *Journal of Economic Theory*, 9(1):43–52.

Zermelo, E. (1929). Die Berechnung der Turnier-Ergebnisse als ein Maximumproblem der Wahrscheinlichkeitsrechnung. *Mathematische Zeitschrift*, 29:436–460.